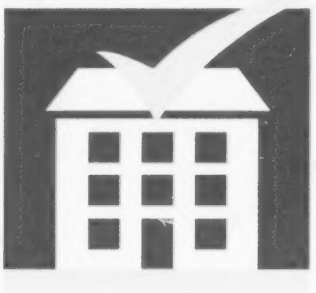
Updated June 26, 2013 to include Language of work

Catalogue no. 99-010-X2011007
ISBN: 978-1-100-22199-1

Reference Guide

Languages Reference Guide



National Household Survey, 2011

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at infostats@statcan.gc.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following toll-free numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-877-287-4369

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

To access this product

This product, Catalogue no. 99-010-X2011007, is available free in electronic format. To obtain a single issue, visit our website, www.statcan.gc.ca, and browse by "Key resource" > "Publications."

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, this agency has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published at www.statcan.gc.ca under "About us" > "The agency" > "Providing services to Canadians."

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2013

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement (www.statcan.gc.ca/reference/licence-eng.html).

Cette publication est aussi disponible en français.

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard symbols

The following symbols are used in Statistics Canada publications:

.	not available for any reference period
..	not available for a specific reference period
...	not applicable
0	true zero or a value rounded to zero
0^s	value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
p	preliminary
r	revised
x	suppressed to meet the confidentiality requirements of the *Statistics Act*
E	use with caution
F	too unreliable to be published
*	significantly different from reference category ($p < 0.05$)

Catalogue no. 99-010-X2011007

Table of contents

Page

Definitions and concepts 4

Classifications 4

Questions 4

Data and other products 5

Data quality 5

Survey process 6

Certification of final estimates 6

Non-response bias 6

Data quality indicators 6

Comparability with other data sources 8

Definitions and concepts

The 2011 National Household Survey (NHS) is an important source of data on knowledge of languages and language use in Canada. The following variables, as defined in the *National Household Survey Dictionary*, Catalogue no. 99-000-X, have been created with language data collected during the NHS on May 10, 2011:

- Knowledge of official languages
- Knowledge of non-official languages
- Home language
- Mother tongue
- First official language spoken
- Language of work

In Canada, 'official languages' refer to English and French. 'Non-official languages' refer to all other languages.

Classifications

Data from language questions in the National Household Survey (NHS) are used to derive summary and detailed variables which provide a linguistic portrait of the population living in Canada. Information is provided on English- and French-speaking communities as well as other language groups, including those who speak Aboriginal languages.

Tables accessible from the Data and other products section of this document show the specific language variables used in data products for the 2011 NHS. As well, a graphic depiction of the 2011 languages classification for Knowledge of non-official languages, Home language, Mother tongue and Language of work is available in the List of figures 1.7 and 1.7A to 1.7F. Appendix 1.3 of the *National Household Survey Dictionary*, Catalogue no. 99-000-X presents the classification for Mother tongue and Home language variables, while Appendix 1.4 presents the classification for the variable Knowledge of non-official languages and Appendix 2.4 presents the classifications for the variable Language of work.

The 2011 NHS language classification is the same as that used in the 2011 Census. The language classification used to disseminate data for the Census of Population and the National Household Survey has been developed over time and corresponds to the ISO 639-3 standard inventory of language identifiers administered by SIL International, as included in the *Ethnologue: Languages of the World* (www.ethnologue.com).

Questions

Most respondents to the 2011 National Household Survey (NHS) received the 2011 National Household Survey Form N1 questionnaire, while respondents living on Indian reserves, in Indian settlements and in other remote areas received the 2011 National Household Survey Form N2 questionnaire. On both questionnaires, data on languages were collected in the sociocultural information section with questions 13, 14, 15(a), 15(b) and 16, and in the labour market activities section with questions 49(a) and 49(b).

- Knowledge of official languages: Question 13
- Knowledge of non-official languages: Question 14
- Home language: Questions 15(a) and 15(b)
- Mother tongue: Question 16
- First official language spoken: Derived from Questions 13, 16 and 15(a)
- Language of work: Questions 49(a) and 49(b).

On the English version of the NHS questionnaire, mark-in circles for English appear first, while on the French version of all NHS forms, the mark-in circles for French appear first.

To assist respondents whose first language was neither English nor French, NHS questions were translated into 31 other languages, including 11 Aboriginal languages.

The N2 language questions were the same as the N1 language questions, but the examples, where provided for write-in responses, were for Aboriginal languages specifically.

More information on the wording and format of the 2011 NHS language questions and the instructions which were provided to respondents for those questions can be found in the 2011 NHS questionnaire and NHS guide and the *National Household Survey Dictionary*, Catalogue no. 99-000-X, entries for Knowledge of official languages, Knowledge of non-official languages, Home language, Mother tongue, First official language spoken and Language of work.

Data and other products

Data on the 2011 National Household Survey (NHS) Knowledge of official languages, Knowledge of non-official languages, Home language, Mother tongue, and First official language spoken variables were released on May 8, 2013 as part of an integrated release with other ethnocultural and Aboriginal variables. As to the data on the Language of work variable, they were released on June 26, 2013 as part on an integrated release with education and labour variables.

The products published using 2011 NHS language data include:

- Data products
  - NHS Profile
  - NHS Data Tables
- Analytical products
- Reference products

For more information on and access to 2011 NHS data, please refer to the 2011 National Household Survey (NHS) website.

Data on language topics for the year 2011 are also available from the 2011 Census. For information on and access to 2011 Census data, please refer to the 2011 Census website. For more information on the comparability of the 2011 National Household Survey with other data sources, please refer to the 'Comparability with other data sources' section of this document, below.

Data quality

The National Household Survey (NHS) underwent a thorough data quality assessment similar to what was done for the 2011 Census of population and past censuses. It consisted of an assessment of various data quality indicators (such as response rate), and an evaluation of the overall results, in comparison with other data sources such as Census of Population data.

Based on the results of this exercise, the NHS estimates for language variables at the national level are consistent with, or similar to, estimates and trends from other data sources such as 2011, 2006 and 2001 Census results.

Survey process

Quality indicators were calculated and assessed at each of the key steps of the survey. During the collection and processing of the data, the quality and consistency of the responses provided were assessed as were the non-response rates. The quality of the imputed responses was assessed after the completion of the control and imputation steps.

Certification of final estimates

Once data processing and imputation were completed, the data were weighted to represent the total Canadian population. These weighted data (the final estimates) were then certified to determine if they were coherent and reliable in comparison to other independent data sources. This is the final stage of data validation. The main highlights of this assessment are presented below.

Non-response bias

Non-response bias is a potential source of error for all surveys including the NHS. This issue arises when the characteristics of those who choose to participate in a survey are different than those who refuse. Statistics Canada adapted its collection and estimation procedures in order to mitigate, to the extent possible, the effect of non-response bias. (For more details please refer to the *National Household Survey User Guide*, Catalogue no. 99-001-X.)

Several data sources were used to evaluate the NHS estimates for language variables such as: 2001, 2006 and 2011 censuses, the Longitudinal Immigrant Database (IMDB) and population projections based on microsimulation.

It is impossible to definitively determine how much the NHS may be affected by non-response bias. However, based on information from other data sources, evidence of non-response bias does exist for certain populations and for certain geographic areas.

For example, based on the estimates and trends from the sources mentioned above, evidence suggests that the population who reported having Malayo-Polynesian languages as their mother tongue is overestimated at the national level. The population reported having Romance languages as their mother tongue appears to be underestimated.

Generally, the risk of error increases for lower levels of geography and for smaller populations. At the same time, the data sources used to evaluate these results are also less reliable making it difficult to certify these smaller estimates.

For more information on NHS non-response bias and mitigation strategies employed by Statistics Canada, please refer to the *National Household Survey User Guide*, Catalogue no. 99-001-X.

Data quality indicators

Of all the quality indicators used for the evaluation, two are presented: the global non-response rate and the imputation rate by question.

- The global non-response rate combines the non-response at the household level and the non-response at the question level. It is provided for geographic areas. The global non-response rate is the key criterion that determines whether or not the NHS results will be released for a given geographic area. Information on the global non-response rate is available in the *National Household Survey User Guide*, Catalogue no. 99-001-X.

- The imputation rate is the proportion of respondents who did not answer a given question or whose response is deemed invalid and for which a value was imputed. Imputation improves data quality by reducing the gaps caused by non-response.

Overall, the imputation rates for NHS language variables were lower than those of the 2011 Census. For the National Household Survey, at the national level, it was 0.6% for mother tongue, ranging from 0.3% in Northwest Territories to 1.8% in Nunavut. For home language, it was 0.6% for Canada, ranging from 0.3% in Northwest Territories to 1.3% in Nunavut. For knowledge of official languages, the imputation rate was 0.7% at the national level, ranging from 0.3% in Northwest Territories to 1.0% in Yukon. These compare against the imputation rates for the 2011 Census: Mother tongue (2.2%), Home language (1.9%) and Knowledge of official languages (1.6%).

For the Language of work variable, the imputation rate was similar to the other variables in Labour market activities section of the questionnaire. At the national level, in 2011, the imputation rate of the Language of work variable was 12.9% and ranged from 3.0% in the Northwest Territories to 14.1% in Ontario. Analysis of the data did not reveal any significant alteration to the language of work data due to imputation. For the data collected with the 2006 Census Form 2B questionnaire, the imputation rate for the Language of work was 4.2% and ranged from 3.2% in Prince Edward Island and Nova Scotia to 7.8% in Nunavut.

Table 1 Imputation rates in the National Household Survey by variable, Canada, provinces and territories

Provinces and territories	Mother tongue	Home language	Knowledge of official languages	Language of work
	percentage			
Canada	**0.6**	**0.6**	**0.7**	**12.9**
Newfoundland and Labrador	0.7	0.7	0.8	13.0
Prince Edward Island	0.8	0.7	0.9	12.3
Nova Scotia	0.7	0.7	0.8	11.1
New Brunswick	0.5	0.5	0.7	11.6
Quebec	0.6	0.6	0.7	11.4
Ontario	0.6	0.6	0.7	14.1
Manitoba	0.6	0.6	0.7	11.7
Saskatchewan	0.6	0.6	0.7	10.9
Alberta	0.5	0.4	0.6	13.2
British Columbia	0.6	0.5	0.7	13.6
Yukon	0.7	0.7	1.0	10.5
Northwest Territories	0.3	0.3	0.3	3.0
Nunavut	1.8	1.3	0.5	4.6

Cross-classification of language variables

Language is often crossed with other variables in a table to analyse a subject in more depth. Data users should be aware that when examining small populations, either by selecting small geographical areas or by crossing multiple variables, the NHS estimates will tend to have greater variability due to sampling error.

Further references related to data quality

For general information on the overall content, collection, design, processing and data quality for the NHS as well as factors that may impact the quality of the NHS data, such as response errors and processing errors, please refer to the *National Household Survey User Guide*, Catalogue no. 99-001-X.

Comparability with other data sources

Statistics Canada disseminates a wealth of data on languages. In addition to disseminating data on languages from the National Household Survey, Statistics Canada publishes language data collected by the Census of Population, the Aboriginal Peoples Survey – Education and Employment, the General Social Survey, the Labour Force Survey and other household surveys.

Many factors affect comparisons of language data across these sources. Amongst other factors, comparability is affected by differences in survey target populations, reference period, sampling and collection methods; question wording, questionnaire format, examples and instructions; approaches to data processing; and the social and political climate at the time of data collection. For additional information, please see the *National Household Survey User Guide*, Catalogue no. 99-001-X.